WEIGHTED WEAK-TYPE INEQUALITY FOR MARTINGALES

ADAM OSĘKOWSKI

Abstract. Let $X = (X_t)_{t\geq 0}$ be a bounded martingale and let $Y = (Y_t)_{t\geq 0}$ be differentially subordinate to X. We prove that if $1 \leq p < \infty$ and $W = (W_t)_{t\geq 0}$ is an A_p weight of characteristic $[W]_{A_p}$, then

$$||Y||_{L^{p,\infty}(W)} \leq C_p[W]_{A_p}||X||_{L^\infty(W)}.$$

The linear dependence on $[W]_{A_p}$ is shown to be the best possible. The proof exploits a weighted exponential bound which is of independent interest. As an application, a related estimate for the Haar system is established.

1. Introduction

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, filtered by $(\mathcal{F}_t)_{t\geq 0}$, a nondecreasing right-continuous family of sub-σ-fields of $\mathcal{F}$, such that $\mathcal{F}_0$ contains all events of probability 0. Let $X = (X_t)_{t\geq 0}$, $Y = (Y_t)_{t\geq 0}$ be adapted, uniformly integrable martingales taking values in $\mathbb{R}^\nu$, $\nu \geq 1$. We also impose the usual regularity assumptions on the paths of these processes, i.e., we assume that X and Y possess right-continuous trajectories that have limits from the left. Next, we denote by $X^* = \sup_{s\geq 0} |X_s|$ the maximal function of X. The symbol $[X, X]$ will stand for the square bracket of X: see e.g. Dellacherie and Meyer [3] for the definition in the case when X is real-valued, and extend to the above vector setting by the formula $[X, X]_t = \sum_{n=1}^{\nu}[X^n, X^n]_t$, where X^n is the n-th coordinate of X. Following Wang [6] and Bañuelos and Wang [1], we say that Y is differentially subordinate to X, if the process $([X, X]_t - [Y, Y]_t)_{t\geq 0}$ is almost surely nonnegative and nondecreasing as a function of t.

The differential subordination implies many interesting martingale inequalities; consult the monograph [5] for almost up-to-date exposition of results in this direction. In [6], Wang proved that if X is bounded almost surely by 1 and Y is differentially subordinate to X, then we have the estimate

$$\mathbb{P}(Y^* \geq \lambda) \leq C(\lambda) := \begin{cases} 1 & \text{if } 0 < \lambda \leq 1, \\ \lambda^{-2} & \text{if } 1 < \lambda \leq 2, \\ e^{2-\lambda}/4 & \text{if } \lambda > 2. \end{cases} \tag{1.1}$$

Furthermore, for each $\lambda > 0$ the constant cannot be improved. In particular, this implies the weak-type bound

$$||Y^*||_{L^{p,\infty}} \leq K_p||X||_{L^\infty}, \qquad 1 \leq p < \infty, \tag{1.2}$$

2010 *Mathematics Subject Classification.* Primary: 60G44. Secondary: 60G42.

Key words and phrases. Martingale, weight, weak-type.

Research supported by Narodowe Centrum Nauki (Poland), grant no. DEC-2014/14/E/ST1/00532.

with the optimal constant equal to

$$K_p = \begin{cases} 1 & \text{if } 1 \le p < 2, \\ \left(p^p e^{2-p}/4\right)^{1/p} & \text{if } p \ge 2. \end{cases}$$

Here, as usual, the weak p-th norm is given by $||\xi||_{L^{p,\infty}} = \sup_{\lambda>0}\left[\lambda^p \mathbb{P}(|\xi| \ge \lambda)\right]^{1/p}$. The estimate (1.1) was obtained with the use of certain special functions constructed by Burkholder in [2]. More precisely, it was shown that for each $\lambda > 0$ there is a function $U_\lambda : \mathbb{R}^\nu \times \mathbb{R}^\nu \to \mathbb{R}$ satisfying the following conditions:

(i) We have $1_{\{|y|\ge\lambda\}} \le U_\lambda(x,y) \le 1$.

(ii) For any $\mathbb{R}^\nu$-valued martingales X, Y such that X is bounded by 1 and Y is differentially subordinate to X, the process $(U_\lambda(X_t,Y_t))_{t\ge0}$ is a supermartingale with $U_\lambda(X_0,Y_0) \le C(\lambda)$ almost surely.

The purpose of this paper is to study weighted versions of the inequalities (1.1) and (1.2). Assume that $W = (W_t)_{t\ge0}$ is a positive, continuous-path and uniformly integrable martingale of mean 1; this process will be called a weight. It defines a new probability measure on $(\Omega, \mathcal{F})$ by $W(A) := \mathbb{E}W1_A$. Let $1 < p < \infty$ be a fixed parameter. Following Izumisawa and Kazamaki [4], we say that W satisfies Muckenhoupt's condition A_p, if

$$[W]_{A_p} := \sup_\tau \left|\left|\mathbb{E}\left[\left\{W_\tau/W_\infty\right\}^{1/(p-1)}\big|\mathcal{F}_\tau\right]^{p-1}\right|\right|_\infty < \infty,$$

where the supremum is taken over the class of all adapted stopping times τ. There are also versions of this condition for $p = 1$: W is an A_1 weight if there is a constant c such that $W^* \le cW$ almost surely; the least c with this property is denoted by $[W]_{A_1}$.

We will establish the following result.

Theorem 1.1. *Suppose that X, Y are $\mathbb{R}^\nu$-valued martingales such that X is bounded by 1 and Y is differentially subordinate to X. Then for any $1 \le p < \infty$ and any A_p weight W we have the estimate*

$$W(Y^* \ge 1) \le 4C(\lambda)^{1/(6[W]_{A_p})}, \qquad \lambda > 0. \tag{1.3}$$

As a consequence, we get the following weak-type bound.

Theorem 1.2. *Suppose that X, Y are $\mathbb{R}^\nu$-valued martingales such that Y is differentially subordinate to X. Then for any $1 \le p < \infty$ and any A_p weight W we have the estimate*

$$||Y^*||_{L^{p,\infty}(W)} \le c_p[W]_{A_p}||X||_{L^\infty(W)}, \tag{1.4}$$

where $c_p = 6pe^{-1}(4e)^{1/p}$. The linear dependence on the characteristic $[W]_{A_p}$ is optimal for each p.

As an application, we will deduce the corresponding weak-type estimate for the Haar system. Let $h = (h_n)_{n\ge0}$ be the family of functions given by $h_0 = \chi_{[0,1)}$, $h_1 = \chi_{[0,1/2)} - \chi_{[1/2,1)}$, and if $n > 1$, then $h_n(t) = h_1(2^k t - \ell)$ where $n = 2^k + \ell$. Given a weight w (i.e., a positive, integrable function with integral equal to 1) on $[0,1)$ and $1 < p < \infty$, we say that w belongs to the (dyadic) class A_p, if

$$[w]_{A_p} := \sup \left(\frac{1}{|I|}\int_I w\mathrm{d}s\right)\left(\frac{1}{|I|}\int_I w^{1/(1-p)}\mathrm{d}s\right)^{p-1} < \infty,$$

where the supremum is taken over the family of all dyadic subintervals of $[0, 1)$ (that is, all intervals of the form $[k2^{-n}, (k+1)2^{-n})$, where $k \in \{0, 1, 2, \ldots, n-1\}$ and $n = 0, 1, 2, \ldots$). Furthermore, w is a (dyadic) A_1 weight, if there is a finite constant $c \geq 1$ such that $Mw \leq cw$ almost everywhere; here M is the dyadic maximal operator, defined by

$$Mw(x) = \sup \frac{1}{|I|} \int_I w \mathrm{d}s$$

and the supremum is taken over all dyadic subintervals of $[0, 1)$ containing x. The smallest constant c with the above property is called the A_1 characteristic of w and is denoted by $[w]_{A_1}$.

We will prove the following statement.

Theorem 1.3. *Let $a_0, a_1, a_2, \ldots, b_0, b_1, b_2, \ldots$ be arbitrary sequences of elements of $\mathbb{R}^\nu$ such that $|a_n| \geq |b_n|$ for all n. Then for any $1 \leq p < \infty$ and any A_p weight w we have*

$$\left\| M\left(\sum_{n=0}^{\infty} b_n h_n\right)\right\|_{L^{p,\infty}(w)} \leq \kappa_p [w]_{A_p} \left\|\sum_{n=0}^{\infty} a_n h_n\right\|_{L^\infty(w)}, \tag{1.5}$$

where κ_p depends only on p. The linear dependence on the A_p characteristic is optimal for each p.

The main result of this paper is the exponential bound (1.3). It will be proved with the use of Burkholder's method (sometimes called in the literature the Bellman function method): we will construct a certain special function of three variables and deduce the exponential bound from the size and concavity properties of this function. This is done in the next section; we also establish the estimate (1.4) there. The final part is devoted to the study in the context of Haar functions.

2. On inequalities (1.3) and (1.4)

It is convenient to split the material into two parts.

2.1. **A special function and its properties.** Let $c \geq 1$ and $1 < p < \infty$ be fixed numbers. Introduce the parameters $a = 3/4$, $\alpha = 1 - 1/(2c)$, $\beta = 1/(6c)$ and consider the domain

$$\mathcal{D}_{p,c} = \{(w, v, z) \in \mathbb{R}_+^3 : 1 \leq wv^{p-1} \leq c\}.$$

Define $B = B_{p,c} : \mathcal{D}_{p,c} \to \mathbb{R}$ by the formula

$$B(w, v, z) = \frac{\left(wv^{p-1} - a\right)^\alpha}{v^{p-1}} z^\beta.$$

We will need the following properties of this object.

Lemma 2.1. *For any $(w, v, z) \in \mathcal{D}_{p,c}$ we have*

$$\frac{1}{4} wz^\beta \leq B(w, v, z) \leq wz^\beta. \tag{2.1}$$

Proof. We must show that

$$\frac{1}{4} \leq \frac{(wv^{p-1} - a)^\alpha}{wv^{p-1}} \leq 1.$$

Observe that the function $t \mapsto (t-a)^\alpha/t$ is increasing: indeed, we have

$$\left(\frac{(t-a)^\alpha}{t}\right)' = \frac{(t-a)^{\alpha-1}((\alpha-1)t+a)}{t^2} \geq 0.$$

Therefore, it is enough to check that $1/4 \leq (1-a)^\alpha$ and $(c-a)^\alpha/c \leq 1$. The first estimate is clear, since $1-a = 1/4$ and $\alpha \in (0,1)$. To show the second, we consider two cases: if $c-a \geq 1$, then $(c-a)^\alpha \leq c-a \leq c$; if $c-a \leq 1$, then $(c-a)^\alpha \leq 1 \leq c$. This completes the proof. □

The key property of B is given in the next statement.

Lemma 2.2. *The Hessian matrix of $-B$ is nonnegative-definite on $\mathcal{D}_{p,c}$. (That is, the function $-B$ is a locally convex function).*

Proof. For brevity, set $\varphi(t) = (t-a)^\alpha$ for $t \geq a$; we will also write $t = wv^{p-1}$ to shorten the notation. The proof rests on Sylvester's criterion. First, note that $B_{ww}(w,v,z) = v^{p-1}\varphi''(t)z^\beta$ is negative, because φ is concave. Next, since

$$B_{wv}(w,v,z) = (p-1)wv^{p-2}\varphi''(t)z^\beta$$

and

$$\begin{aligned} B_{vv}(w,v,z) = & p(p-1)v^{-p-1}\varphi(t)z^\beta - p(p-1)wv^{-2}\varphi'(t)z^\beta \\ & + (p-1)^2w^2v^{p-3}\varphi''(t)z^\beta, \end{aligned}$$

we derive that

$$\det\begin{bmatrix} B_{ww} & B_{wv} \\ B_{vw} & B_{vv} \end{bmatrix} = p(p-1)v^{-2}\big[\varphi(t) - t\varphi'(t)\big]\varphi''(t)z^{2\beta}.$$

However, $\varphi(t) - t\varphi'(t) = (t-a)^{\alpha-1}(t(1-\alpha)-a)$ is negative when $t \leq c$; this shows that the above determinant is positive (since $(w,v) \in \mathcal{D}_{p,c}$). It remains to show that the determinant of the full Hessian is nonpositive:

$$\det\begin{bmatrix} B_{ww} & B_{wv} & B_{wz} \\ B_{vw} & B_{vv} & B_{vz} \\ B_{zw} & B_{zv} & B_{zz} \end{bmatrix} \leq 0.$$

Add to the second column the first column multiplied by $-(p-1)w/v$; then add to the second row the first row multiplied by $-(p-1)w/v$. Then the above inequality amounts to saying that the determinant

$$\det\begin{bmatrix} v^{p-1}\varphi''(t)z^\beta & 0 & \beta\varphi'(t)z^{\beta-1} \\ 0 & p(p-1)v^{-p-1}(\varphi(t)-t\varphi'(t))z^\beta & -\beta(p-1)v^{-p}\varphi(t)z^{\beta-1} \\ \beta\varphi'(t)z^{\beta-1} & -\beta(p-1)v^{-p}\varphi(t)z^{\beta-1} & \beta(\beta-1)v^{-p+1}\varphi(t)z^{\beta-2} \end{bmatrix}$$

is nonpositive. It is easy to see that the powers of z and v appearing above do not affect the sign of the determinant; in other words, we must show that

$$\begin{aligned} \det & \begin{bmatrix} \varphi''(t) & 0 & \beta\varphi'(t) \\ 0 & p(p-1)(\varphi(t)-t\varphi'(t)) & -\beta(p-1)\varphi(t) \\ \beta\varphi'(t) & -\beta(p-1)\varphi(t) & \beta(\beta-1)\varphi(t) \end{bmatrix} \\ & = -(\beta-1)p(t\varphi'(t)-\varphi(t))\varphi(t)\varphi''(t) + \beta p(\varphi'(t))^2(t\varphi'(t)-\varphi(t)) \\ & \quad - \beta(p-1)\varphi^2(t)\varphi''(t) \leq 0, \end{aligned}$$

or, after some manipulations,

$$(\beta-1)(1-\alpha)((\alpha-1)t+a)+\beta\alpha((\alpha-1)t+a)+\beta\frac{p-1}{p}(1-\alpha)(t-a)\leq 0.$$

It is easy to see that it suffices to show the bound for $p\to\infty$ and $t=c$; then the estimate is the strongest and reads

$$\beta\leq\frac{(1-\alpha)((\alpha-1)c+a)}{\alpha a}.$$

Plugging the values of α, β and a prescribed at the beginning, we get the desired assertion. □

2.2. **Proof of** (1.3) **and** (1.4)**.** Any A_1 weight automatically belongs to all A_p classes, $p>1$, and we have $[W]_{A_p}\leq[W]_{A_1}$. Thus we may assume that $p>1$ in our considerations below. We will use the following useful interpretation of A_p weights. Fix such a weight W and let $c=[W]_{A_p}$. Furthermore, let $V=(V_t)_{t\geq 0}$ be the martingale given by $V_t=\mathbb{E}(W_\infty^{1/(1-p)}|\mathcal{F}_t)$, $t\geq 0$. Note that Jensen's inequality implies $W_\tau V_\tau^{p-1}\geq 1$ almost surely; furthermore, the A_p condition is equivalent to the reverse bound

$$W_\tau V_\tau^{p-1}\leq c \qquad \text{with probability 1.}$$

In other words, an A_p weight of characteristic equal to c gives rise to a two-dimensional martingale (W,V) taking values in the domain $\mathcal{D}_{p,c}$. In addition, this martingale terminates at the lower boundary of this domain: $W_\infty V_\infty^{p-1}=1$ almost surely. A nice feature is that this is a full characterization: given any martingale pair (W,V) (with continuous-path W of mean 1) taking values in $\mathcal{D}_{p,c}$ and terminating at the set $wv^{p-1}=1$, one easily checks that its first coordinate is an A_p weight with $[W]_{A_p}\leq c$.

We are ready for the proof of the main estimate (1.3). Let X, Y, W be martingales as in the statement of Theorem 1.1 and, given $\lambda>0$, let $U_\lambda:\mathbb{R}^\nu\times\mathbb{R}^\nu\to\mathbb{R}$ be the special function of Burkholder [2], with the properties listed in the introductory section. Then the process $Z_t=U_\lambda(X_t,Y_t)$ is a supermartingale; let $Z=Z_0+M+A$ be the Doob-Meyer decomposition for Z (cf. [3]). Let us also consider the auxiliary process $\xi_t=(W_t,V_t,Z_t)$, $t\geq 0$, where V is given as above, and let $c=[W]_{A_p}$. The function $B=B_{p,c}$ is of class C^∞ (more precisely, it extends to a C^∞ function on some open set containing $\mathcal{D}_{p,c}$), so we are allowed to apply Itô's formula to obtain

$$B(\xi_t)=I_0+I_1+I_2+I_3/2+I_4,$$

where

$$\begin{aligned}
I_0&=B(\xi_0),\\
I_1&=\int_0^t B_w(\xi_{s-})\mathrm{d}W_s+\int_0^t B_v(\xi_{s-})\mathrm{d}V_s+\int_0^t B_z(\xi_{s-})\mathrm{d}M_s,\\
I_2&=\int_0^t B_z(\xi_{s-})\mathrm{d}A_s,\\
I_3&=\int_0^t D^2B(\xi_{s-})\mathrm{d}[W,V^c,Z^c]_s,\\
I_4&=\sum_{0<s\leq t}\Big[B(\xi_s)-B(\xi_{s-})-B_v(\xi_{s-})\Delta V_s-B_z(\xi_{s-})\Delta Z_s\Big],
\end{aligned}$$

where I_3 is the abbreviated form of the sum of all the second-order terms. Note that in I_4 there are no terms $-B_w(\xi_{s-})\Delta W_s$, since the weight W is assumed to have continuous paths. Let us analyze the summands $I_0 - I_4$. By the right inequality in (2.1), we have

$$I_0 \leq W_0 Z_0^\beta = W_0 U_\lambda(X_0, Y_0)^\beta \leq W_0 C(\lambda)^\beta.$$

The stochastic integrals in I_1 have expectation zero. The process A coming from the Doob-Meyer decomposition is nonincreasing and $B_z \geq 0$, so the term I_2 is nonpositive. We also have $I_3 \leq 0$, which follows directly from Lemma 2.2 and a standard approximation of the integrals by Riemann-type sums (see e.g. [6] for a similar reasoning). Finally, each summand appearing in I_4 is nonpositive, which is the consequence of concavity of B inside its domain. Putting all the above facts together, we obtain $\mathbb{E}B(\xi_t) \leq C(\lambda)^\beta \mathbb{E}W$, which combined with the left inequality from (2.1) gives

$$W(|Y_t| \geq \lambda) = \mathbb{E}W_t 1_{\{|Y_t|\geq\lambda\}} \leq \mathbb{E}W_t U(X_t, Y_t)^\beta \leq 4\mathbb{E}B(\xi_t) \leq 4C(\lambda)^\beta \mathbb{E}W.$$

To pass from Y to Y^*, we exploit a well-known stopping time argument. Fix $\varepsilon \in (0, \lambda)$ and let $\tau = \inf\{t : |Y_t| \geq \lambda - \varepsilon\}$. Since $\{Y^* \geq \lambda\} \subseteq \{\tau < \infty\}$, we may write

$$W(Y^* \geq \lambda) \leq \lim_{t\to\infty} W(|Y_{\tau\wedge t}| \geq \lambda - \varepsilon) \leq 4C(\lambda - \varepsilon)^\beta \mathbb{E}W_0.$$

We have $\mathbb{E}W_0 = 1$, by the very definition of a weight. Letting $\varepsilon \to 0$ and using the fact that the function $\lambda \mapsto C(\lambda)$ is continuous, we get the desired exponential estimate (1.3).

Now the proof of (1.4) is straightforward. By homogeneity, we may assume that $||X||_{L^\infty(W)} = 1$. Then we use (1.3) and the elementary estimate $C(\lambda) \leq e^{1-\lambda} \leq e^{1/\beta-\lambda}$ to get

$$\lambda^p w(Y^* \geq 1) \leq \lambda^p C(\lambda)^\beta \cdot 4\mathbb{E}W \leq \lambda^p e^{-\lambda\beta} \cdot 4e.$$

Optimizing the right-hand side over λ, we obtain the weak-type inequality (1.4).

3. Inequalities for the Haar system

3.1. **Proof of** (1.5)**.** As in the probabilistic context, we may and do assume that p is strictly larger than 1. Fix two sequences $(a_n)_{n\geq0}$, $(b_n)_{n\geq0}$ as in the statement of Theorem 1.3. We will embed the functions $f = \sum_{n=0}^\infty a_n h_n$, $g = \sum_{n=0}^\infty b_n h_n$, $w = \sum_{n=0}^\infty c_n h_n$ and $w^{1/(1-p)} = \sum_{n=0}^\infty d_n h_n$ into certain continuous-time martingales satisfying differential subordination. To this end, first we rewrite the formulas for f and g in terms of the Rademacher sequence $r_1 = h_1$, $r_2 = h_2+h_3$, $r_3 = h_4+h_5+h_6+h_7$, ... Let $(\mathcal{G}_n)_{n\geq1}$ be the filtration generated by $(r_n)_{n\geq1}$. Then there are $(\mathcal{G}_n)$-predictable sequences $(\bar{a}_n)_{n\geq1}$, $(\bar{b}_n)_{n\geq1}$, $(\bar{c}_n)_{n\geq1}$ and $(\bar{d}_n)_{n\geq0}$ (the first two of which take values in $\mathbb{R}^\nu$) such that $|\bar{b}_n| \leq |\bar{a}_n|$ almost surely and $f = a_0 + \sum_{n=1}^\infty \bar{a}_n r_n$, $g = b_0 + \sum_{n=1}^\infty \bar{b}_n r_n$, $w = c_0 + \sum_{n=1}^\infty \bar{c}_n h_n$ and $w^{1/(1-p)} = d_0 + \sum_{n=1}^\infty \bar{d}_n h_n$. In particular, the predictability implies that for each n, the variables $\bar{a}_n$, $\bar{b}_n$, $\bar{c}_n$ and $\bar{d}_n$ are functions of r_1, r_2,..., r_{n-1}:

$$\bar{a}_n = \bar{a}_n(r_1, r_2, \ldots, r_{n-1}), \qquad \bar{b}_n = \bar{b}_n(r_1, r_2, \ldots, r_{n-1})$$

and similarly for $\bar{c}_n$ and $\bar{d}_n$. Now let $(B_t)_{t\geq0}$ be a standard Brownian motion starting from 0 and let $(\tau_n)_{n\geq0}$ be a sequence of stopping times of B given inductively by $\tau_0 \equiv 0$ and

$$\tau_{n+1} = \inf\{t > \tau_n : |B_t - B_{\tau_n}| = 1\}.$$

Then $(B_{\tau_{n+1}} - B_{\tau_n})_{n\geq 0}$ is a sequence of independent Rademacher variables, so has the same distribution as the sequence $(r_n)_{n\geq 1}$ considered above. Define the processes $X = (X_t)_{t\geq 0}$, $Y = (Y_t)_{t\geq 0}$, $W = (W_t)_{t\geq 0}$ and $V = (V_t)_{t\geq 0}$ by the formulas

$$X_t = a_0 + \sum_{k=1}^{\infty} \bar{a}_n(B_{\tau_1} - B_{\tau_0}, \ldots, B_{\tau_{n-1}} - B_{\tau_{n-2}})(B_{\tau_n\wedge t} - B_{\tau_{n-1}\wedge t}),$$

$$Y_t = b_0 + \sum_{k=1}^{\infty} \bar{b}_n(B_{\tau_1} - B_{\tau_0}, \ldots, B_{\tau_{n-1}} - B_{\tau_{n-2}})(B_{\tau_n\wedge t} - B_{\tau_{n-1}\wedge t}),$$

$$W_t = c_0 + \sum_{k=1}^{\infty} \bar{c}_n(B_{\tau_1} - B_{\tau_0}, \ldots, B_{\tau_{n-1}} - B_{\tau_{n-2}})(B_{\tau_n\wedge t} - B_{\tau_{n-1}\wedge t}),$$

$$V_t = d_0 + \sum_{k=1}^{\infty} \bar{d}_n(B_{\tau_1} - B_{\tau_0}, \ldots, B_{\tau_{n-1}} - B_{\tau_{n-2}})(B_{\tau_n\wedge t} - B_{\tau_{n-1}\wedge t}).$$

Then Y is differentially subordinate to X (which follows directly from the assumption $|b_n| \leq |a_n|$ for each n). Furthermore, the pair (W, V) terminates at the set $\{(x, y) : xy^{p-1} = 1\}$ (since the pair $(w, w^{1/(1-p)})$ takes its values there). Now we will show that

$$\text{(3.1)} \quad \text{the pair } (W, V) \text{ takes values in } \{(x, y) : 1 \leq xy^{p-1} \leq \max\{2^{p-1}, 2\}[w]_{A_p}\},$$

which will imply the A_p property of W. To check this, observe that the distribution of (W_{τ_n}, V_{τ_n}) is the same as that of $(\sum_{k=0}^n \bar{c}_k r_k, \sum_{k=0}^n \bar{d}_k r_k) = \mathbb{E}((w, w^{1/(1-p)})|\mathcal{G}_n)$ and hence, by the A_p property of w, is concentrated on $\{(x, y) \in \mathbb{R}^2_+ : 1 \leq xy^{p-1} \leq [w]_{A_p}\}$. Let us look at the behavior of the pair (W, V) on the interval $[\tau_n, \tau_{n+1}]$ for some fixed n. Suppose that $(W_{\tau_n}, V_{\tau_n}) = (x, y)$; then $(W_{\tau_{n+1}}, V_{\tau_{n+1}}) \in \{(x_+, y_+), (x_-, y_-)\}$, where $1 \leq x_\pm y_\pm^{p-1} \leq [w]_{A_p}$ and $(x_- + x_+)/2 = x$, $(y_- + y_+)/2 = y$. Furthermore, on the interval $[\tau_n, \tau_{n+1}]$, the pair (W, V) moves along the line segment joining (x_-, y_-) and (x_+, y_+). Therefore, to show (3.1), it is enough to establish the following statement.

Lemma 3.1. *Assume that $c > 1$ and suppose that points P, Q and $R = (P+Q)/2$ lie in the set $\{(x, y) : 1 \leq xy^{p-1} \leq c\}$. Then the whole line segment PQ is contained within $\{(x, y) : 1 \leq xy^{p-1} \leq \max\{2^{p-1}, 2\}c\}$.*

Proof. Using a simple geometrical argument, it is enough to consider the case when the points P and R lie on the curve $wv^{p-1} = c$ (the upper boundary of $\{(x, y) : 1 \leq xy^{p-1} \leq c\}$) and Q lies on the curve $wv^{p-1} = 1$ (the lower boundary of the set). Then the line segment RQ is contained within $\{(x, y) : 1 \leq xy^{p-1} \leq c\}$, and hence also within $\{(x, y) : 1 \leq xy^{p-1} \leq \max\{2^{p-1}, 2\}c\}$, so it is enough to ensure that the segment PR is contained in $\{(x, y) : 1 \leq xy^{p-1} \leq \max\{2^{p-1}, 2\}c\}$. Let $P = (P_x, P_y)$, $Q = (Q_x, Q_y)$ and $R = (R_x, R_y)$. We consider two cases. If $P_x < R_x$, then

$$P_y = 2R_y - Q_y < 2R_y,$$

so the segment PR is contained in the quadrant $\{(x, y) : x \leq R_x,\ y \leq 2R_y\}$. Consequently, PR lies below the hyperbola $xy^{p-1} = 2^{p-1}c$ passing through $(R_x, 2R_y)$; this proves the assertion in the case $P_x < R_x$. In the case $P_x \geq R_x$ the reasoning is similar: then the line segment PR lies below the hyperbola $xy^{p-1} = 2c$ passing through $(2R_x, R_y)$. □

Proof of (1.5). We know that W is an A_p weight and Y is differentially subordinate to X, so (1.4) gives

$$W(Y^*\geq 1)\leq c_p^p[W]_{A_p}^p||X||_{L^\infty(W)}^p.$$

It follows from the above construction that for each n, $(X_{\tau_n}, Y_{\tau_n}, W_{\tau_n}, V_{\tau_n})$ has the same distribution as the quadruple $(\sum_{k=0}^n a_kh_k, \sum_{k=0}^n b_kh_k, \sum_{k=0}^n c_kh_k, \sum_{k=0}^n d_kh_k)$ and, in particular, $\sup_{n\geq 0}|Y_{\tau_n}|$ has the same distribution as Mg. Furthermore, by (3.1), we have $[W]_{A_p}\leq \max\{2^{p-1},2\}[w]_{A_p}$, so the above weak-type bound implies

$$w(Mg\geq 1)\leq c_p^p\max\{2^{p-1},2\}^p[w]_{A_p}^p||f||_{L^\infty(w)},$$

which is precisely the claim. □

3.2. **On the linear dependence on the characteristic.** Now we will show that the linear dependence in the weak-type bound is optimal in the context of Haar system with real-valued coefficients; this will automatically show that this dependence is optimal in the probabilistic setting as well. Consider the functions

$$f=\frac{1}{3}+\frac{2}{3}\sum_{n=0}^\infty(-1)^{n+1}h_{2^n},\qquad g=\frac{1}{3}+\frac{2}{3}\sum_{n=0}^\infty h_{2^n}$$

and introduce the weight

$$w=1+\left(1-\frac{1}{c}\right)\sum_{n=0}^\infty\left(2-\frac{1}{c}\right)^n h_{2^n}.$$

It is easy to check that $f=\sum_{n=0}^\infty(-1)^n\chi_{[2^{-n-1},2^{-n})}$ and hence f is bounded in absolute value by 1. On the other hand, on the set $[2^{-n-1},2^{-n})$ we have $h_1=1$, $h_2=1$, $\ldots$, $h_{2^{n-1}}=1$ and $h_{2^n}=-1$, so $g=\frac{1}{3}+\frac{2}{3}(n-1)$ there and

$$\left\{g\geq \frac{1}{3}+\frac{2}{3}(n-1)\right\}=[0,2^{-n}). \tag{3.2}$$

Concerning w, we see that on $[2^{-n-1},2^{-n})$ we have

$$w=1+\left(1-\frac{1}{c}\right)\left[1+\left(2-\frac{1}{c}\right)+\ldots+\left(2-\frac{1}{c}\right)^{n-1}-\left(2-\frac{1}{c}\right)^n\right]=\frac{1}{c}\left(2-\frac{1}{c}\right)^n,$$

so in particular w is positive (and hence is a weight). Furthermore, w is a nonincreasing function on $[0,1)$, so its maximal function can be computed as follows. If $x\in[0,1)$ and k is the unique positive integer such that $x\in[2^{-k-1},2^{-k})$, then

$$\begin{aligned}Mw(x)&=\frac{1}{|[0,2^{-k})|}\int_{[0,2^{-k})}w\mathrm{d}s\\&=2^k\sum_{n=k}^\infty\int_{[2^{-n-1},2^{-n})}w\mathrm{d}s\\&=2^k\sum_{n=k}^\infty 2^{-n-1}\cdot\frac{1}{c}\left(2-\frac{1}{c}\right)^n=\left(2-\frac{1}{c}\right)^k.\end{aligned}$$

Consequently, we have $Mw=cw$ on $[0,1)$ and hence w is an A_1 weight with $[w]_{A_1}=c$. We obviously have $||f||_{L^\infty(w)}=1$ and, by (3.2),

$$w\left(g\geq\frac{1}{3}+\frac{2}{3}(n-1)\right)=\int_{[0,2^{-n})}w\mathrm{d}s=2^{-n}\left(2-\frac{1}{c}\right)^n=\left(1-\frac{1}{2c}\right)^n.$$

Now take $n = \lfloor c \rfloor + 2$ and $\lambda = \frac{1}{3} + \frac{2}{3}(n-1) \geq \frac{2}{3}c$. Then

$$||g||^p_{L^{p,\infty}(w)} \geq \lambda^p w(g \geq \lambda) \geq \left(\frac{2}{3}c\right)^p \left(1 - \frac{1}{2c}\right)^{\lfloor c \rfloor + 2} \geq \kappa_p c^p ||f||^p_{L^\infty(w)},$$

for some constant κ_p depending only on p. This proves that the linear dependence is indeed optimal.

Acknowledgments

The author would like to thank an anonymous referee for the careful reading of the first version of the paper and several helpful comments and suggestions.

References

[1] R. Bañuelos and G. Wang, Sharp inequalities for martingales with applications to the Beurling-Ahlfors and Riesz transformations, Duke Math. J. **80** (1995), 575–600.

[2] Burkholder, D. L. (1991). Explorations in martingale theory and its applications, École d'Été de Probabilités de Saint-Flour XIX—1989, 1–66, Lecture Notes in Math., 1464, Springer, Berlin.

[3] C. Dellacherie and P.-A. Meyer, Probabilities and potential B: Theory of martingales, North Holland, Amsterdam, 1982.

[4] M. Izumisawa and N. Kazamaki, Weighted norm inequalities for martingales, Tohoku Math. J. **29** (1977), 115–124.

[5] A. Osękowski, Sharp martingale and semimartingale inequalities, Monografie Matematyczne **72** (2012), Birkhäuser, 462 pp.

[6] G. Wang, Differential subordination and strong differential subordination for continuous time martingales and related sharp inequalities, Ann. Probab. **23** (1995), 522–551.

Department of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland

E-mail address: ados@mimuw.edu.pl